NASA TECHNICAL REPORT

NASA TR R-472 c.1

NASA TR R-472

A PROCEDURE FOR DETECTION AND MEASUREMENT OF INTERFACES IN REMOTELY ACQUIRED DATA USING A DIGITAL COMPUTER

Kenneth H. Faller

Lyndon B. Johnson Space Center
Houston, Texas 77058

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JULY 1977

1. Report No.	2. Government Accession No.	3. Recipien
NASA TR R-472		

4. Title and Subtitle
A PROCEDURE FOR DETECTION AND MEASUREMENT OF INTERFACES IN REMOTELY ACQUIRED DATA USING A DIGITAL COMPUTER

5. Report [
July 1977

6. Performing Organization Code
JSC-11562

7. Author(s)
Kenneth H. Faller

8. Performing Organization Report No.
JSC S-467

9. Performing Organization Name and Address
Lyndon B. Johnson Space Center
Houston, Texas 77058

10. Work Unit No.
177-55-89-00-72

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
National Aeronautics and Space Administration
Washington, D.C. 20546

13. Type of Report and Period Covered
Technical Report

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract
A technique developed and evaluated for the detection and measurement of surface feature interfaces in remotely acquired data is described. A computer implementation of this technique has been effected to automatically process categorized data derived from various sources such as the Landsat multispectral scanner and other scanner-type sensors. The basic elements of the operational theory of the technique are described together with details of the procedure. An example application of the technique to the analysis of tidal shoreline length is given with a breakdown of manpower requirements. The technique as applied to Landsat multispectral scanner data is evaluated with respect to accuracy and precision.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Remote sensors Landsat satellite Interface boundary analysis Shoreline analysis and measurement Remote data acquisition and processing Computer processing of satellite data	STAR Subject Category: 43 (Earth Resources)

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	32	$4.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

NASA — JSC

JSC Form 1424 (Rev Jul 74)

CONTENTS

TABLES

FIGURES

A PROCEDURE FOR DETECTION AND MEASUREMENT OF INTERFACES IN REMOTELY ACQUIRED DATA USING A DIGITAL COMPUTER

By Kenneth H. Faller*
Lyndon B. Johnson Space Center

SUMMARY

A technique for the accurate detection and measurement of surface feature interfaces in remotely acquired data has been developed and evaluated. The technique has been implemented on a digital computer to automatically process categorized data derived from various sources such as the Landsat multispectral scanner and other scanner-type sensors.

The computer-implemented interface detection and measurement technique is in the form of a 10-step procedure. Any type of imagery that has been categorized into two or more classes and is in a sampled, properly formatted form may be used in the interface analysis. Typically, multispectral scanner data that have been processed by a pattern recognition technique are used to generate an image classified into categories that define the interface feature to be detected and measured.

The technique was successfully applied to measurement of the Alabama tidal shoreline. Landsat multispectral scanner imagery was used in the analysis together with an existing simple two-category classifier. Approximately 100 man-hours were required to complete the demonstration project.

Application of the technique to multiple Landsat data sets has established the precision of the technique as 3.5 percent, and comparison with measurements made using traditional methodology indicates that the Landsat-based measurement agrees with measurements made on 1:24 000-scale maps to better than 5 percent.

INTRODUCTION

The linear measurement of the interface between two features is potentially very valuable, especially when coupled with measurement of the areal extent of the delineated features. One important measurement of this type that can be made is of land, water, and shoreline, and it was as a shoreline analysis tool that the technique to be described was developed.

*Earth Resources Laboratory, Slidell, Louisiana 70458.

Interfaces are very important in the physics, chemistry, biology, sociology, and politics of the world. For example, thermal energy is transferred across the boundary between a powerplant cooling-water discharge and ambient water; erosion takes place at the shoreline; it is across the land/water interface in the marsh that nutrients flow, as the inorganic materials for primary producers and the organic materials for higher marine life; at the interface between the farmland and small creek or stream, fertilizers and insecticides applied to the fields become agricultural runoff; it is at the interface of the highway construction project with the swamp, wilderness, or tundra that the environmental impact of the highway is first felt; the extent of interfaces between land cover types is one factor wildlife managers use in evaluating wildlife habitats; it is at the interface with industrial development that residential neighborhood quality frequently declines; it is at the flood-plain boundary that the hazard to development suddenly changes; and limits of ownership of land and mineral rights are often determined by particular land/water boundaries.

Given an understanding of the mechanisms involved in environmental dynamics, a definition of the interfaces involved becomes integral to a quantitative analysis of natural and human-induced processes. Properly directed management of Earth resources requires this type of analysis, and one important input to the analysis will come from monitoring, detecting, and measuring the interface involved. Because many of the natural boundaries are very extensive, and are also readily detectable in remotely acquired data, automatic analyses of aircraft and satellite data are particularly attractive as sources of information for environmental management.

The technique described in this document is designed to provide an accurate and inexpensive means of extracting information about interfaces from remotely acquired imagery. Actual use of the technique on satellite (Landsat) and aircraft data to detect and measure the shoreline in marsh areas demonstrates one application of interest to geographers and resource managers. Only the basic elements of the interface measurement theory are presented here. This report contains descriptions of data that are appropriate for interface analysis, of the steps in the procedure for extracting the interface measurement from the data, and of the products generated in the analysis. An example of an actual interface measurement problem is reported. A breakdown of manpower requirements is given for this example. Accuracy, repeatability, and effects of sampling unit size were studied in the analysis of a series of data sets.

The computer-implemented interface detection and measurement technique was originally suggested by G. C. Thomann. The algorithm for performing this task was implemented on the computer by J. G. Glydewell of Lockheed Electronics Co., Inc. Glydewell was also responsible for performing the computer work necessary for the geometric corrections and analyses performed in verifying the technique. Jon Brown, also of Lockheed Electronics, was the technician who performed most of the work described in the demonstration section of this report. The photographic work required for the tidal effects study was performed by Charles Morgan, formerly of Lockheed Electronics. C. D. Sapp of the Geological Survey of Alabama represented that agency and provided the information and direction required to enable generation of Alabama tidal shoreline analysis desired by that State government.

As an aid to the reader, where necessary the original units of measure have been converted to the equivalent value in the Système International d' Unités (SI). The SI units are written first, and the original units are written parenthetically thereafter.

DATA

Any kind of imagery that has been classified into two or more categories and is in a sampled, properly formatted form may be used in the interface analysis. The typical problem will involve the use of multispectral scanner (MSS) data which have been processed by some pattern recognition technique to generate an image classified into categories that define the interface feature to be detected and measured. The data are not restricted to this type of imagery. For example, thermal scanner imagery can be analyzed to generate an image showing water of normal temperature and a thermal effluent plume, and the interface analysis technique can be used to measure the interface at the surface of the water between the plume and the ambient water for dispersion studies.

Data in the form of maps or photographs can also be manipulated. This type of data may be sampled and digitized to produce a map or an image in a format suitable for processing by means of the interface analysis technique.

A sampled image is in the form of a series of small picture elements, often called pixels. The image may be thought of as a matrix with rows corresponding to the scan lines generated by the scanner, with the individual matrix elements being the pixels. The picture element has dimensions determined by the scanner and the platform; the separation of the center points of the pixels within a row of the matrix determines the "width" of the element, and the center-to-center distance between pixels in a column of the matrix determines the "height" of the element. Figure 1 illustrates the relationship of the sampled imagery to the scene portrayed.

THEORY

The description of the interface analysis technique theory of operation presented herein is not intended to be exhaustive or fully comprehensive.

Interface Detection and Measurement

Two basic problems must be solved in the analysis of an interface in classified data. The interface first must be detected, and then it must be measured. These problems are treated separately in this report but are handled as one by the computer in the interest of maximum efficiency.

Just as the image consists of individual picture elements, the interface defined in the analysis is composed of discrete line segments, know as interface elements. Interface is a linear phenomenon, whereas the picture elements

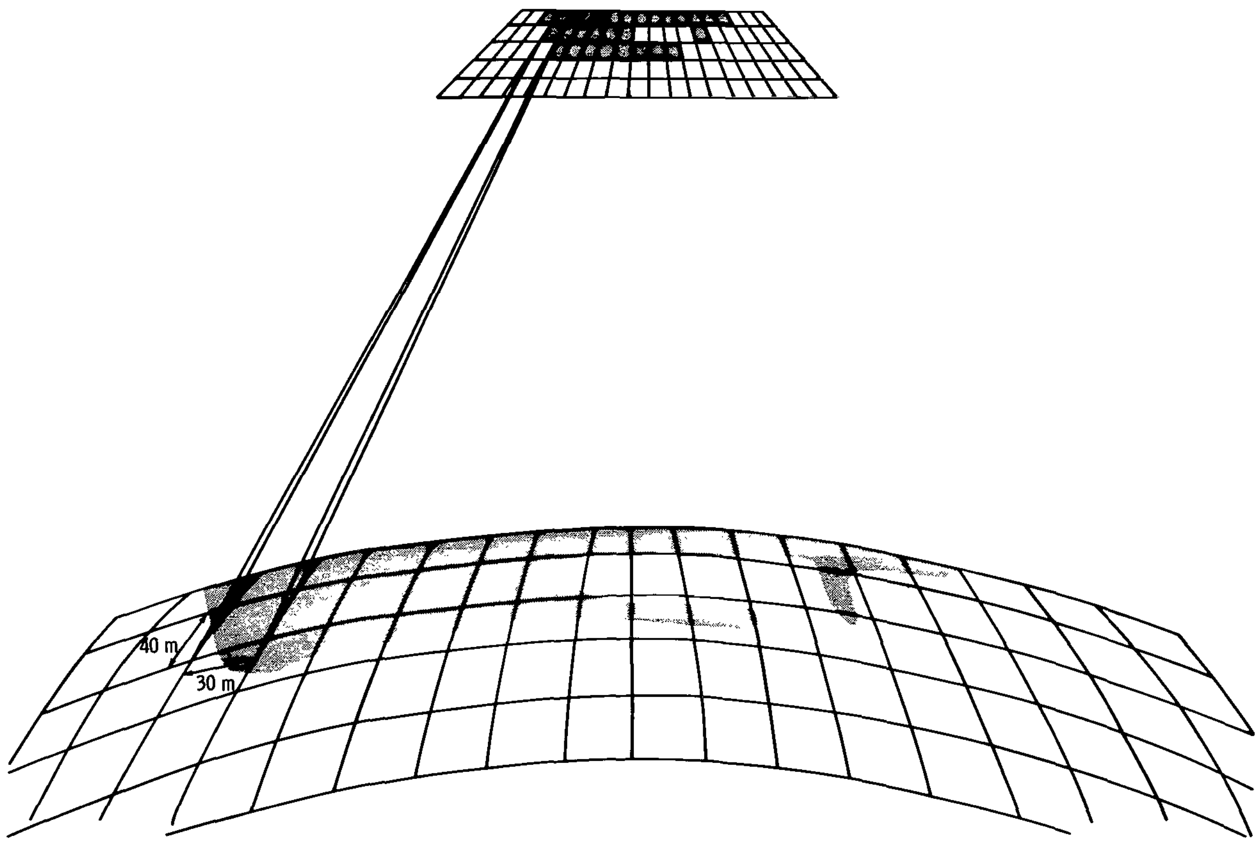


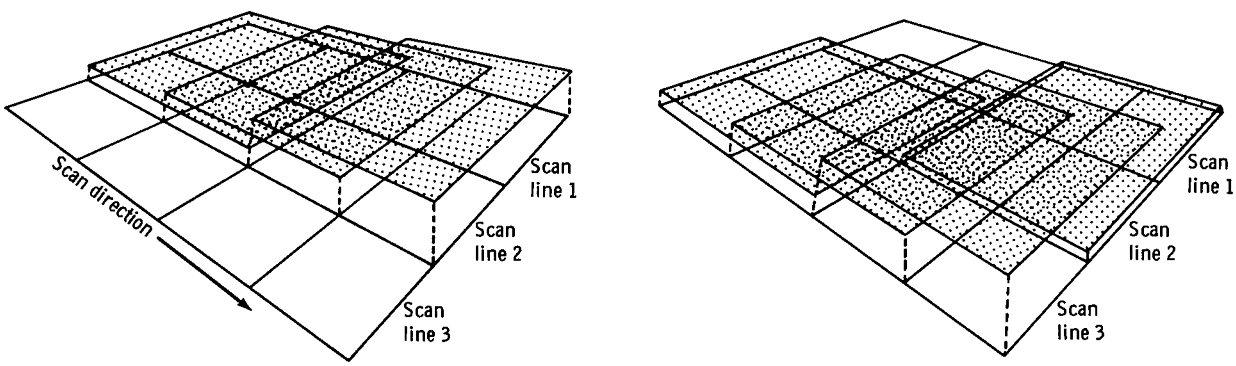
Figure 1.- Scanner image concept. The scanner image (top) corresponds to the Earth scene (bottom). Each element in the scanner data represents the average conditions of the area on the surface within the instantaneous field of view, which is 30 by 40 meters in this illustration.

are areal in nature. Consequently, the interface element is defined as the common edge of two picture elements that represent different classes. Therefore, a single picture element may have as many as four interface elements associated with it.

As developed, the interface detection algorithm uses only two classes to define the interface. A two-category product is not frequently generated by standard pattern recognition analyses; therefore, the many classes which may be present in the data must be grouped into superclasses. For example, if a measurement of the land/water interface in a swamp is desired, 10 classes representing individual vegetation species or groups of species may be required to define "land" and two classes representing lake water and river water may define "water." The 10 vegetation classes would then be grouped into the land superclass, and the water classes would be grouped to give the water superclass. An additional vegetation class may represent floating vegetation; therefore, in an analysis of shoreline length, one would group the floating vegetation with the water superclass. The remainder of this report will not be concerned with this grouping; instead, the classified data will be considered to consist of only two categories, although these may be superclasses such as those just described. Areas which fall into classes included in neither superclass are excluded from the analysis and do not contribute to area or interface determination.

The actual detection of an interface between the two classes is accomplished by examining the entire data set sequentially through a two-pixel-wide by two-pixel-high examination window. If two classes are present within the window, there must be an interface; if only one of the two possible classes is present, there is no interface. The window is scanned over the data, as illustrated in figure 2. The window is shifted to the right one element at a time so that each pixel appears first at the lower right corner of the window, then at the lower left corner. After completing a scan across one row of the image array, the window is shifted down one element to the next row of the matrix and scanned again. The elements that previously appeared in the bottom row of the window then appear in the top row. Each element is passed over four times for a complete analysis of the data.

The contribution of an interface element to the total interface measurement is dependent on the orientation of the element. The first case to consider is that in which interface elements are oriented either parallel (horizontal) or perpendicular (vertical) to the direction of the sensor scan. For these elements, the contribution is the respective dimension of the picture element itself. Two other cases present special problems. First, if a straight interface feature is oriented at an angle that is neither parallel nor perpendicular to the scan direction, the feature will appear as a jagged line, constructed with alternating horizontal and vertical elements. The other problem arises with the sampling of smooth curves, where the finite sampling grid causes a smooth, rounded interface feature to appear as a sharp corner. These two types of interface distortion are illustrated in figure 3.



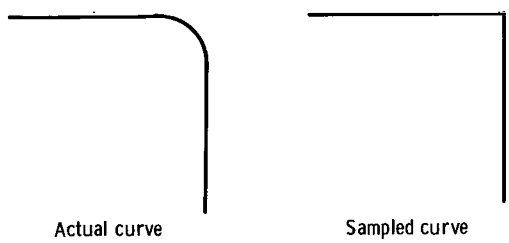
(a) Step 1. (b) Step 2.

Figure 2.- Two-step interface detection procedure in which a four-element examination window (shaded squares) is passed along scan lines 1 and 2, then along scan lines 2 and 3, overlapping one scan line. An interface is detected when two different classes are found within the window.

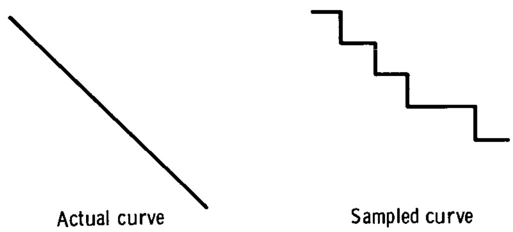
When either diagonal- or corner-type features are detected within the examination window, picture elements adjacent to the window are examined to determine which of the two types is actually present.

The diagonals are subdivided into three types, examples of which are shown in figure 4. The same elements surrounding the examination window used to distinguish the diagonals from the corners also determine the type of diagonal by the slope of the indicated interface feature relative to the sensor scan. The slopes for the three diagonal types are $2V/H$, V/H, and $V/2H$, where V and H are the picture element dimensions perpendicular and parallel to the scan direction, respectively. The contribution to the interface length for each of the diagonal types is $\sqrt{H^2 + (2V)^2}$, $\sqrt{H^2 + V^2}$, and $\sqrt{(2H)^2 + V^2}$, respectively. Because of the manner in which the scanning examination window reveals the interface elements, the presence of the first or third types of diagonal interface element introduces an error into the total length being accumulated by causing an extra element to be detected and counted either along the scan line or perpendicular to it. Thus, the appropriate correction is made when this error occurs.

When examination of the pixels surrounding the basic examination window indicates that a curved surface feature has been squared off by the sampling,



(a) Curved (corner) interface.



(b) Diagonal interface.

Figure 3.- Examples of interface distortion. Problems arise when interface features on the surface are curved, and when they are straight but neither parallel nor perpendicular to scan direction.

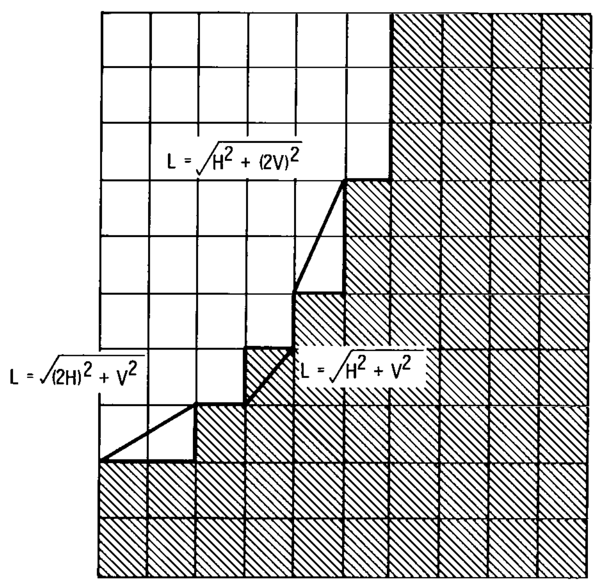


Figure 4.- Diagonal interfaces. Three types of diagonal interface elements are considered in the analysis, each with a unique contribution to the total length L. The crosstrack pixel dimension is H (along scan line), and the other dimension is V.

the contribution of this corner interface element is added to the total being accumulated. This contribution is computed as the sum of one-half of the picture element dimensions and one-fourth of the mean of the perimeters of the ellipse that can be inscribed and circumscribed on the picture element. The 0.5H and 0.5V contributions are accumulated with the sums of the along-scan-line and across-scan-line interfaces. The additional length of the curve is computed from the relation

$$C = 0.60355AE(k)$$

where E is the complete elliptical integral, $k = \sqrt{1 - B^2/A^2}$, and A and B are the long and short pixel dimensions, respectively.

Area Measurement

As the original classified data are being grouped into the superclasses, the number of elements of each superclass is accumulated by counting the picture elements which are being placed in it. Upon completion of the processing of the data, the number of picture elements in each class is multiplied by the area of an individual element to give the total area of each superclass. This method of area mensuration assumes that all geometric rectification has been performed during a preceding phase of processing.

Display Generation

It is usually desirable to generate a display of the analysis results to show the interface that has been detected and measured. This display is accomplished by creating a third class (in addition to the original two superclasses) which represents the interface for display purposes only. When a picture element of the second class is found to be adjacent to an element of the first class, it is replaced by an element of the third, or interface, class. The modified display then shows the original data grouped into the two superclasses with the newly created interface class substituted for elements of the second superclass when they occur along the boundary between the two superclasses.

PROCEDURE

The processing sequence that carries the original data into a display showing the interface between two classes and that lists the interface length and area of each class is basically the same for the various types of data which are available and suitable for this work. The most commonly used scanner data thus far have been provided by Landsat; consequently, the description of the interface analysis procedure is directed toward that type of data. The procedure is divided into 10 basic steps, which are outlined in figure 5.

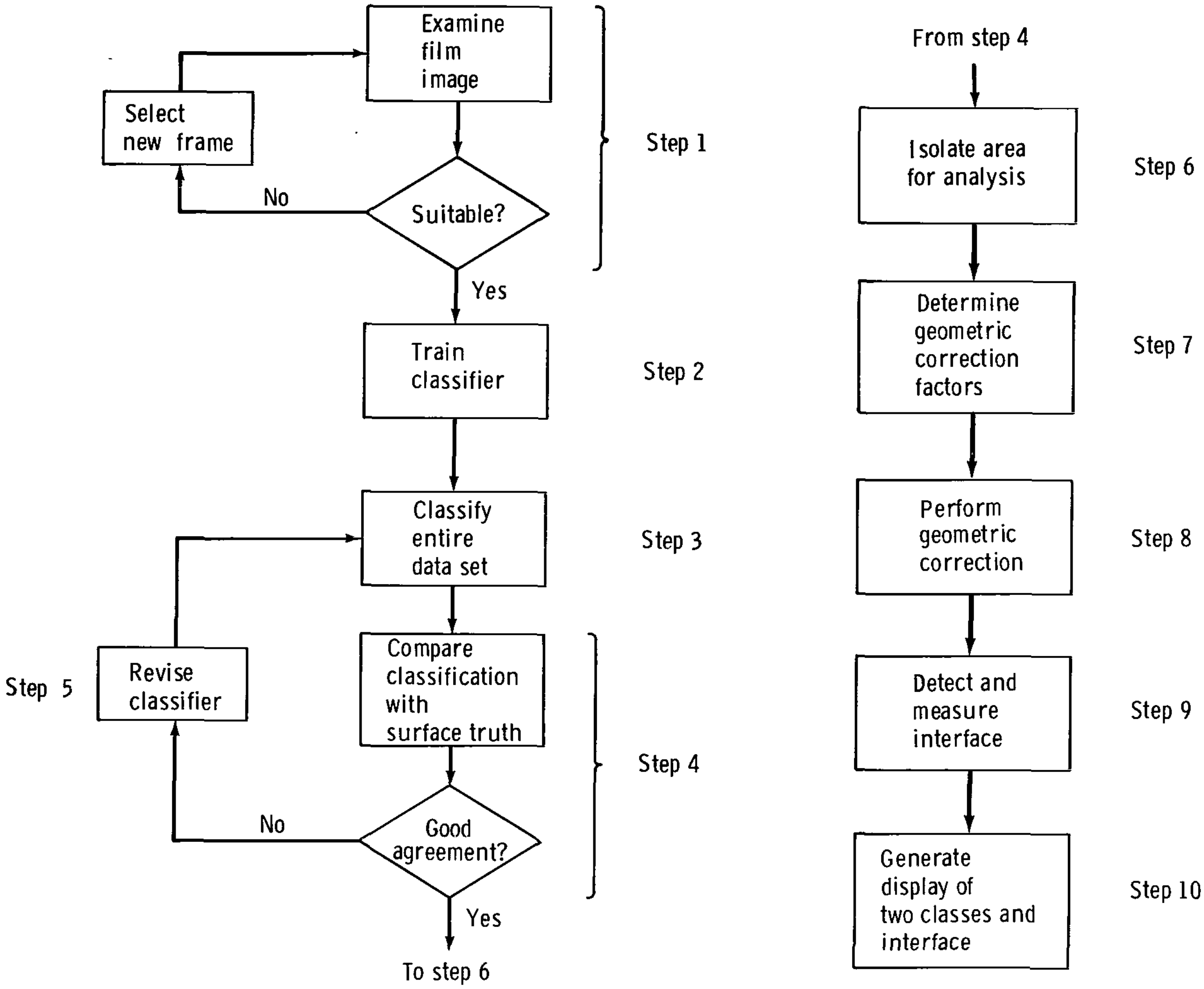


Figure 5.- Interface analysis processing flow chart.

The analysis begins with a determination of the suitability of the data for the proposed task (step 1). The data are initially examined as a hard-copy image or a visual display; the investigator notes factors such as whether the area to be studied is completely included in the data set, whether clouds will interfere with the analysis, and whether the data appear to be of sufficient quality to distinguish the classes required for the interface measurement. If the data set is found to be suitable for the task, the computer-compatible magnetic tapes are obtained.[1] All subsequent work makes use of these tapes.

[1]In the case of Landsat data, this step involves acquiring the tapes from the Department of the Interior Earth Resources Observation Systems Data Center or, in certain cases, from the Department of Agriculture, from the National Oceanic and Atmospheric Administration, or from NASA; in the case of aircraft scanner data, decommutation of the pulse-code-modulated recorded signal from the aircraft tape is required to generate computer-compatible tapes.

Steps 2 to 5 involve classification of the data into the various categories required to define the interface to be analyzed. Typical classification procedures are described in greater detail in other documents (e.g., ref. 1), but the basic elements relevant to this analysis are described in the following paragraphs.

Most classification work is based on spectral pattern recognition. The computer must be "trained" to recognize spectral signatures of the classes to be found in the data (step 2). These signatures are developed by an analysis of spectral data from areas known to be representative of the classes in general. These areas are referred to as training samples since they are used to "train" the computer to recognize the spectral signatures; i.e., the characteristic spectral responses of the target categories. The entire data set is then processed by the classifier, with each picture element being tested against the spectral signatures developed from the training samples and categorized according to a decision rule which is normally built into the pattern recognition algorithm (step 3).

After a classified product has been generated, it must then be evaluated relative to surface truth information (step 4). The product may be compared with aerial photography which has been analyzed by a photointerpreter or with information acquired through field activities. If the product matches the surface truth information with sufficient accuracy (which must be determined by the user for each application), the classification is complete. However, if the number of errors in the classification is unacceptable, the input parameters of the pattern recognition software must be revised (step 5). The classification is repeated with the new parameters and the new product evaluated, and this procedure is repeated for each frame until an acceptable product is generated. The number of repetitions is generally reduced with user experience. If, after extensive analysis of the data, it is determined that the data are not suitable for separating the desired classes, a new set of data may be obtained that will allow the classification to be performed accurately; or, in the worst case, the problem may be modified.

Step 6 in the interface analysis procedure is the isolation of the data upon which the analysis is to be performed. Normally, the data set will cover an area larger than the area of interest. For example, in the demonstration analysis presented later in this document, the problem was measurement of the tidal shoreline. It was therefore necessary to eliminate all of the land/water interface along riverbanks, farm ponds, and so forth where tidal flow was not present. The study area may be defined on a hard-copy display of the classified data for reference in this step, in which all picture elements that are to be excluded from the analysis are reclassified into a unique category which is recognized as having been excluded by the final interface detection/analysis computer program.

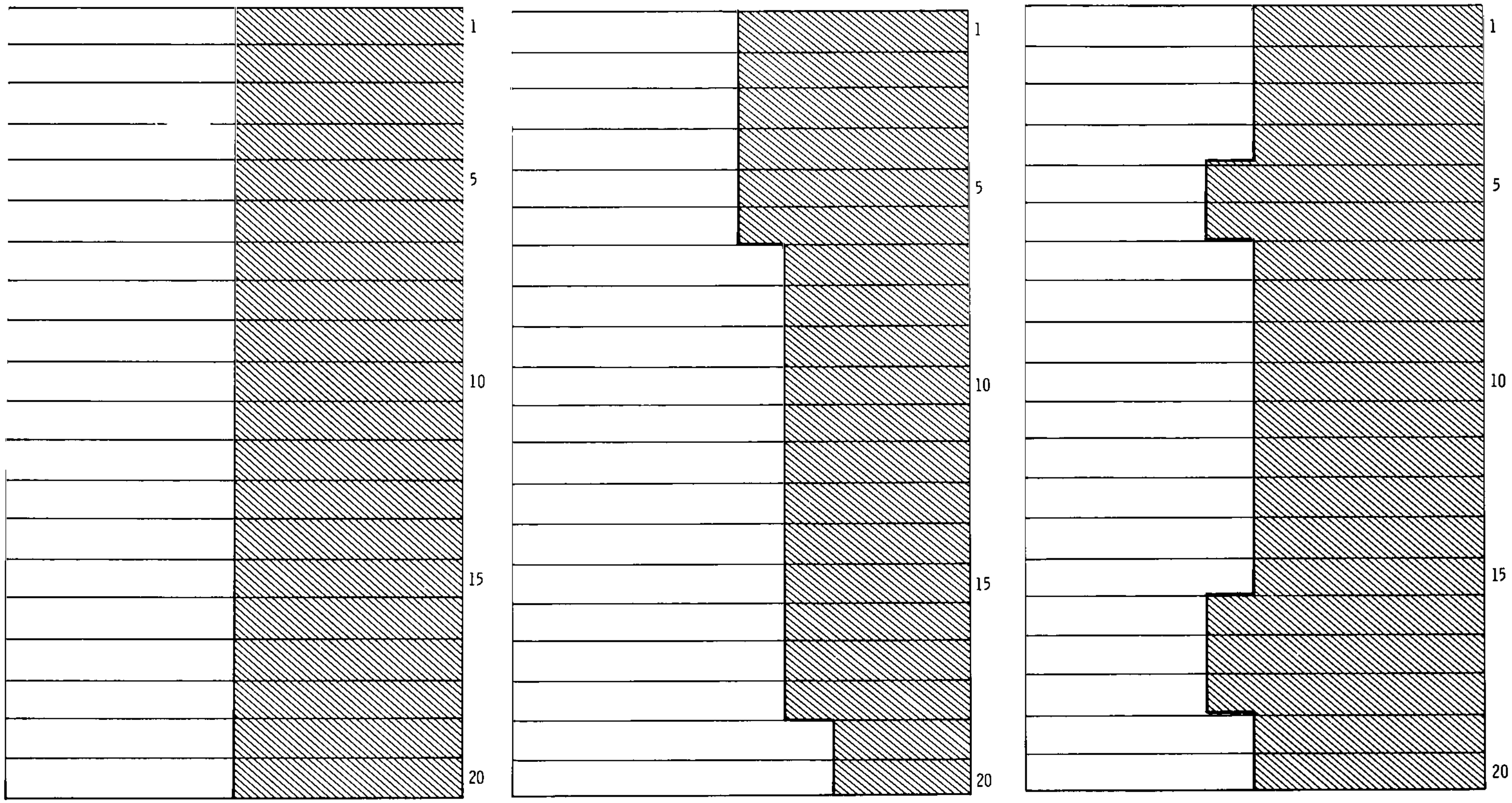
The data generated by the NASA Earth Resources Laboratory (ERL) pattern recognition software are contained on magnetic tape in a particularly simple format. After a preamble containing information such as scan-line number and length, each 6-bit character representing a picture element contains a number in the range of 0 to 63 which corresponds uniquely to a category, with the number 0 reserved for unclassified areas. A unique number may be assigned to excluded areas, but all classes not grouped into either superclass are excluded. The superclasses are internally assigned a value of 1 or 2 by the interface analysis program, whereas excluded areas are internally assigned the value of 3.

Alternatively, a polygon may be defined that includes only the area for which the analysis is to be performed. The vertices of the polygon are defined in terms of scan-line number and number of picture elements from the start of the scan line, and these coordinates are input with the unmodified data set to the interface analysis computer program, which operates only on the pixels within the polygon. The polygon may be any simple, closed curve having no restriction of convexity. The vertices may first be located by drawing the polygon on a standard map. They are then defined in terms of standard map coordinates, which may then be converted to scan-line and element coordinates through geographic referencing of the data set. Direct definition of the polygon is also possible by using either a hard-copy display of the data or an interactive image processing system.

The seventh step of the analysis is the determination of the geographic parameters that relate the scanner data to the actual surface scene. This relation requires geometric correction for distortion inherent in the data and determination of the scale of the data.

The problem of geometric correction is somewhat different when considering the interface detection/measurement analysis than when approaching a straightforward mapping problem. For this measurement, the exact positioning of a given feature is not the primary consideration; the principal consideration is that straight lines remain straight and curved lines retain their original curvature. This is a very difficult problem in working with aircraft data when the roll correction of the scanner fails to fully compensate for the aircraft roll or when the altitude, the heading, or the speed varies. The problem with aircraft data will not be addressed in this document.

The use of satellite data requires consideration of the problem introduced by the rotation of the Earth beneath the satellite. The distortion is seen as an apparent shifting of scan lines to the east for satellites in a descending polar orbit, such as Landsat. The effect of this distortion on the interface measurement is to shorten features oriented in a generally northeast-southwest direction and to lengthen features oriented in a generally northwest-southeast direction. The correction for this skewing of the image depends on the altitude of the satellite, the latitude of the study area, and the scan rate of the sensor. It is typically made by shifting entire scan lines to the west as the satellite proceeds southward along its orbital path. Use of Landsat data requires consideration of a special feature of the multispectral scanner flown on that platform. Because the Landsat MSS collects six scan lines of data simultaneously, this distortion is introduced into the image only at intervals of six scan lines. Skew correction is performed by periodically shifting the data to the west while working south in the image, but no correction should be made within this six-line unit, because its geographic integrity is not disturbed by rotation of the Earth. All corrections for skew in the Landsat data



(a) A feature as it appears on the Earth surface and in properly corrected satellite-borne scanner data.

(b) The feature (fig. 6(a)) as it appears in uncorrected scanner data from a satellite platform.

(c) The feature (fig. 6(a)) as it appears in data corrected without considering the six-scan-line unit.

Figure 6.- Correction for rotation of the Earth beneath a satellite.

must be made between these units, as illustrated in figure 6. If the corrections are made within the six-line unit, linear features transversing the proper shift point, the erroneous shift point, or both will normally be lengthened.

An interface between two features as they might appear on the surface of the Earth and in Landsat data is shown schematically in figure 6. In figure 6(a), the actual feature is shown. In figure 6(b), the feature is shown as it appears in uncorrected Landsat MSS data, as it would be extracted from the computer-compatible tapes on which the data are recorded and provided to the user. Figure 6(c) was constructed by shifting these raw computer-compatible tape (CCT) data one element to the west every 11 scan lines, the proper rate of correction based on a statistical analysis of an entire frame of data. Although the image as a whole will be portrayed accurately with this correction, with no net displacement or rotation of the feature introduced, the feature itself will be distorted. Proper correction would shift the data 1 element to the west every 12 scan lines, with 1 additional westward shift introduced every 132 scan lines, to give 12 shifts in 132 scan lines just as the simpler method does.

The scale of the imagery is determined from the physical dimensions of an individual picture element on the surface of the Earth. When this scale has been established, the contributions of the various interface types can be computed.

The two dimensions of the picture element and the skew correction factor (expressed as the frequency at which additional westward shift of one pixel must be made) are best determined directly from the image, as opposed to theoretical calculations based on sensor and platform characteristics. This determination is made by locating points both on accurate maps and in the imagery, measuring the distance separating the points on the map, and using these distances together with the separation of the points in the image in terms of scan lines and scan elements in a least-squares error analysis. An equation of the form

$$D = \sqrt{\left(S_V \ \Delta V\right)^2 + \left[S_H\left(\Delta H - S_K \ \Delta V\right)\right]^2}$$

is used, where D is the distance between the points, S_V is the vertical scale factor (i.e., the dimension of the picture element perpendicular to the scan direction), ΔV is the number of scan lines between the points in the image, S_H is the horizontal scale factor (the dimension of the picture element along the scan line), ΔH is the number of elements along the scan line separating the points in the image, and S_K is the skew correction factor. Only the two scale factors are used directly in the interface analysis software, but when satellite data are used, the skew correction must be applied first.

It is important that the points be taken in the same portion of the data that will be analyzed because the size of the picture element will vary with scan angle in most kinds of data. The Landsat MSS data contain a small distortion along the scan lines introduced as a result of the nonlinear rate of motion of the oscillating mirror that scans the Earth scene. Aircraft data are usually acquired with such large scan angles that one cannot assume a constant instantaneous field of view of the sensor. This latter problem should be corrected by rectifying the data before the interface analysis is performed. When large areas are to be subjected to the analysis, they should be divided into units for which constant scale factors can be assumed, and the analysis should then be performed on the individual segments and the results combined.

The classified data from Landsat are usually processed at ERL using a computer program known as SKEWCOR, which makes use of the previously computed skew correction parameter to perform the proper shifting of data to compensate for the rotation of the Earth beneath the satellite. This program is to be used only temporarily, and a more general and more accurate geographic correction and reference program is planned for the future.

After being corrected for distortion (processing step 8), the data are ready to be processed by the interface detection and measurement program SHORL (step 9). The program is loaded with the scale factors and with the codes corresponding to the classes to be included in each of the two superclasses. Any classes not included in either superclass are disregarded in the processing and treated as unclassified. The scan-line number with which processing is to begin and the number with which it is to terminate may be read in.

The interface detection and analysis program SHORL then reads the input data tape, and a display of the two classes and the interface is generated on a second magnetic tape with an identical format. The length of the interface between the two classes and the area of each are computed and listed. Figure 7 is an example of the display product, which is generated as the final step of processing.

DEMONSTRATION

Discussion with personnel of the Geological Survey of Alabama revealed the need for an objective measurement of the Alabama shoreline. Several estimates had been made by conventional means, but there were large discrepancies among them. Use of satellite data for measurement of the shoreline, by way of the computer-implemented interface analysis algorithm, provided the opportunity to make the measurement on the basis of defined objective criteria. The resolution of the data is defined by the resolution of the Landsat multispectral scanner, and the geographic limits of the analysis can be explicitly defined. Perhaps most important, the element of human error in the manipulation of the opisometer is completely eliminated; such error is an important consideration in the conventional manual measurement.

This section of the report contains a step-by-step description of the work performed in developing the shoreline analysis for the State of Alabama.

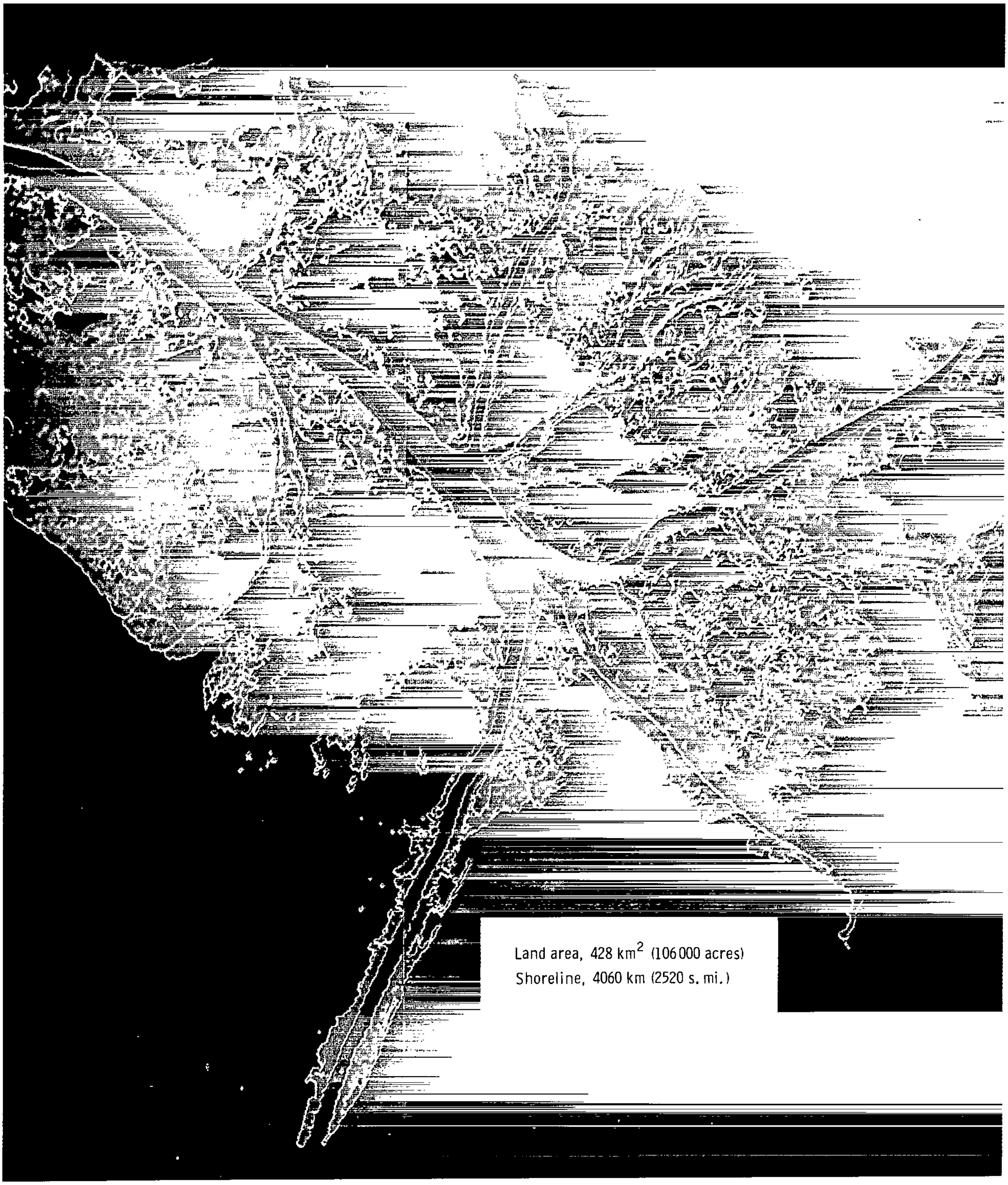
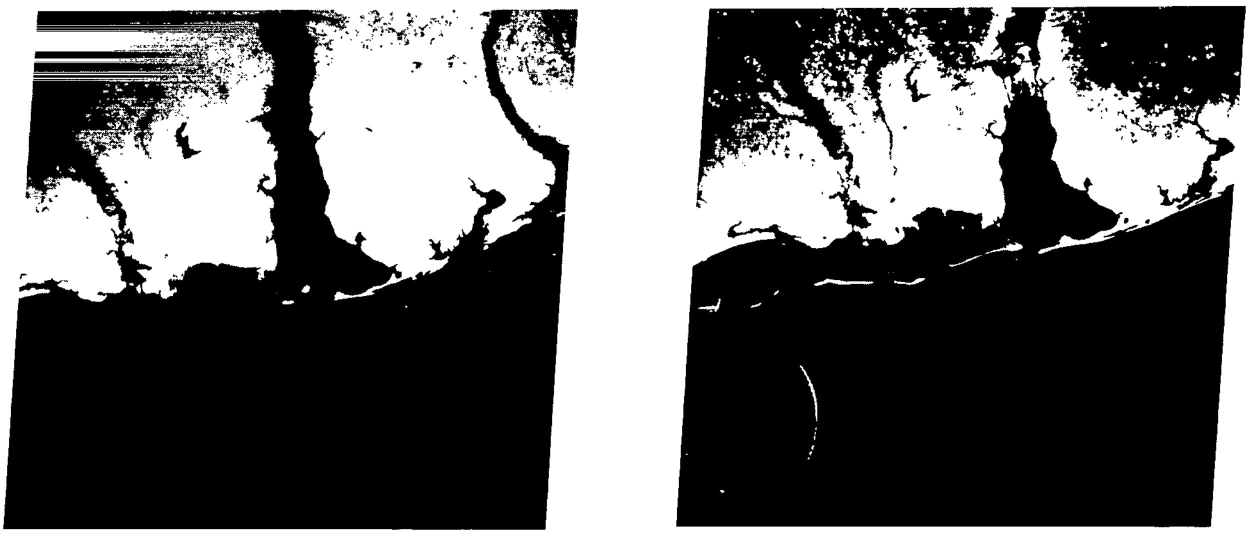


Figure 7.- Display products generated by the interface analysis program using data acquired over the active delta of the Mississippi River on January 16, 1973, by Landsat 1. The yellow area represents the interface which has been detected and measured.

The time required for each step is only approximate, because the personnel involved in the work were assigned to other tasks simultaneously, and some modifications to the procedure and the software were instituted during the period in which the analysis was performed. The procedure has been further streamlined since the completion of this demonstration. The work was performed with a low-cost data analysis system, an interactive image processing system similar to configuration 4 described in reference 2, and a Univac 1108 system. All software is suitable for implementation on a minicomputer system such as specified for configuration 4.

Data Selection (Processing Step 1)

To simplify problems with floating or overhanging vegetation, the decision was made that data from winter periods would be used in the analysis. The area to be studied covered the entire southern portion of Alabama. Because one of the Landsat groundtracks passes through the center of the area, the data are available on a single data frame. The film images covering the area during the desired period were selected, one from December 1972 and the other from December 1973. The two frames are shown in figure 8. Conditions were somewhat different for the two dates, because river discharges were considerably greater in December 1972 than in December 1973. Because small streams are not detected in one set of data but are detected in the other, as swelling of small streams may bring them into the size range that can be resolved by the scanner, analysis of the shoreline under these two conditions should represent two extreme cases. After the data had been selected, CCT's for the two frames were ordered. Tapes 2 to 4 of the four-tape set were required to cover the study area.



(a) December 1972. (b) December 1973.

Figure 8.- Landsat frames for Alabama shoreline study.

The first phase of the analysis required an effort of 4 hours on the part of the investigator.

Classification (Processing Steps 2 to 5)

A simple two-category classifier was used in this demonstration project. The classification technique is a land/water discriminator developed at the NASA Lyndon B. Johnson Space Center for use with Landsat data and modified by ERL. The classifier, known as Water Search (ref. 3) uses the intensity measurements in the green band (0.5 to 0.6 micrometer) and in one of the infrared bands (0.8 to 1.1 micrometers). This infrared band alone provides very good definition of land and water because water strongly absorbs radiation in that spectral region, whereas typical land features strongly reflect this radiation. The spectral signature of water is defined as being all values in the infrared band that are less than a particular value. This particular value is a function of the intensity in the green band. The intensity of green light reflected from the area being classified serves to adjust this infrared intensity value, the decision point, for varying water turbidity levels, a factor which can cause recognition problems in areas typified by muddy river discharges or marshes.

The spectral signature for water was developed as the first phase of the classification. Areas known to be land and areas known to be water were located in the image. Marshland, urban areas, agricultural regions, and forest lands provided the nonwater spectral information, whereas the Gulf of Mexico, the Mississippi Sound, upper and lower Mobile Bay, and inundated marsh areas provided the spectral information on the appearance of water. A graph with axes representing the green band intensity and the infrared band intensity was developed; data points corresponding to land features were differentiated from those corresponding to water features. The line that best separated the two sets of points was then drawn through the data. This line defined the discriminator for the classification of the entire data set. Classification was accomplished by comparing each pixel to the curve and categorizing those elements falling to the right of the curve as land and those falling to the left as water.

The classifier was applied to the areas being used to develop the discriminator, and the results showed excellent agreement with interpretation of aerial photography. Then, the classifier was applied to the complete data set and a classified image was produced. The complete procedure was performed for both data sets (1972 and 1973), since spectral signatures could not be used on data sets other than the one from which they were derived. These images were carefully compared with aerial photography acquired in February 1973 (NASA Earth Resources Aircraft Project, Flight Number 73-023). Considering the fact that river conditions were very different on the two dates, excellent agreement between the two Landsat data sets and the photography was achieved. The classification of the data was a rather lengthy process; the determination of spectral signatures required an effort of approximately 30 hours on the part of a technician and 5 hours on the part of the investigator. The classification of the entire data set required 4 hours by the computer operator. Subsequent improvements made in the Water Search procedure have reduced the time required in this phase of the work by at least a factor of seven.

Isolation (Processing Step 6)

The requirement of the Geological Survey of Alabama was for a measurement of the Alabama tidal shoreline. The development of the analysis within that constraint required that all the land/water interface outside the tidal area within the State be eliminated as well as all interface outside the Alabama State boundaries. Using a display of one of the images resulting from the land/water classification, Geological Survey personnel delineated the area to be analyzed by drawing a line on the image defining the study area (fig. 9).

Using the data analysis station interactive update program, this limit was then transferred to the computer-compatible data. By means of this procedure, all data outside the study area were reduced to a unique value that would be excluded from the interface analysis.

Delineation of the study area required an effort of approximately 3 hours on the part of the investigator and one representative of the Geological Survey. The interactive data manipulation required 18 hours of work by a technician and a computer operator, and evaluation of the intermediate product required an effort of 1 hour by the investigator.

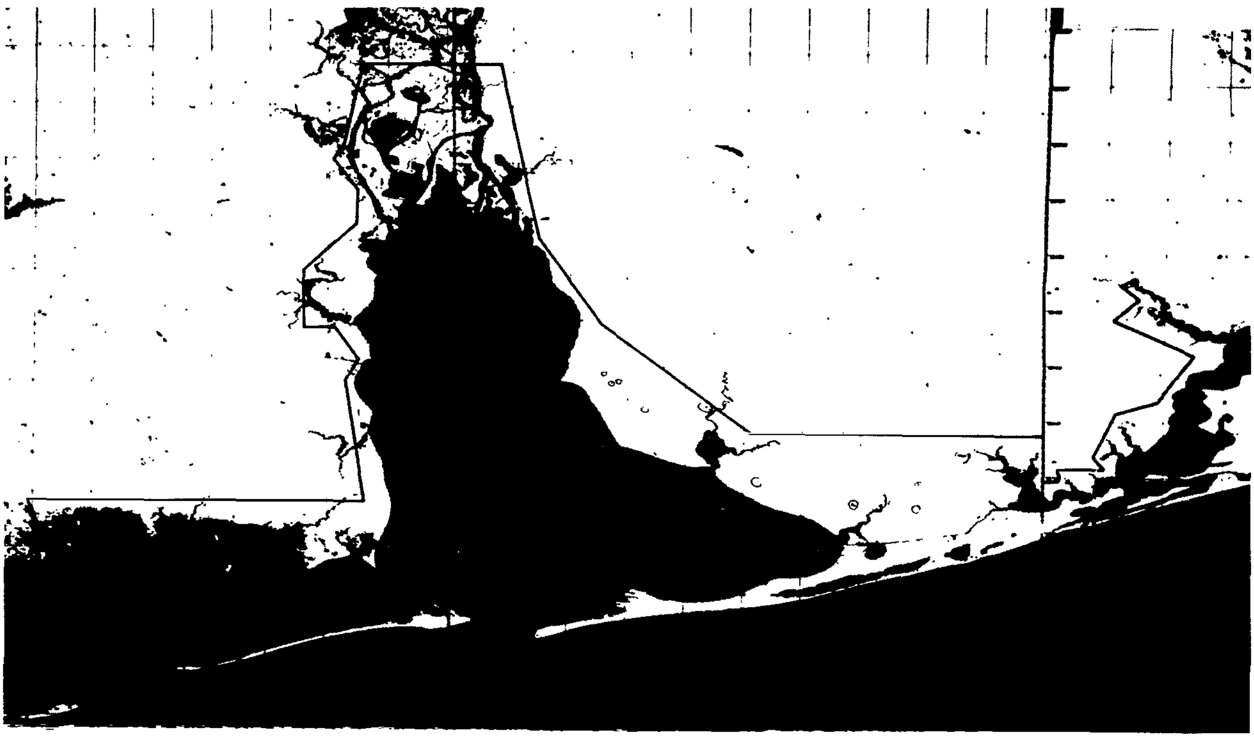


Figure 9.- Delineation by State personnel of area to be analyzed.

Scale Determination (Processing Step 7)

Using the land/water thematic map generated by program Water Search, a set of landmarks was located. The scan-line number and element within the scan were determined for each landmark by using the thematic map, whereas the actual distances separating the points were measured on National Ocean Survey chart 1266. Eight points were identified, and five interpoint distances were measured on the chart. The picture element dimensions then were determined to be 57.34 meters parallel to the scan line and 80.80 meters perpendicular to the scan line for the 1972 data and 57.94 and 80.04 meters, respectively, for the 1973 data.

The determination of the picture element dimensions required 6 hours of effort by a technician.

Geometric Correction and Interface Measurement (Processing Steps 8 to 10)

Program SKEWCOR was used to perform the correction for the rotation of the Earth beneath the satellite. The number of elements to be processed in each scan line was limited at this stage of processing to eliminate the non-video information (calibration data not used in this analysis) appearing at the end of the second and third tapes.

After the skew correction had been performed, the three tapes for each date were processed using the interface analysis computer program. The results of the analysis for all three tapes for each date are summarized in table I. After the analyses had been performed on the complete data set for each date, the interface display tapes generated by the analysis software were

TABLE I.- ALABAMA TIDAL SHORELINE

LENGTH ANALYSIS RESULTS

Date	Tape	Shoreline length, km
1972	2	732.5
1972	3	557.2
1972	4	162.8
Total	--	1452.5
1973	2	136.5
1973	3	795.4
1973	4	230.7
Total	--	1162.6
Mean	--	1307.5

output on an electrostatic printer/plotter. The products are shown in figures 10(a) and 10(b), where the land appears white, the water light gray, and the shoreline black. Only the area included in the study is shown in the final products.

(a) 1972.

(b) 1973.

Figure 10.- Alabama shoreline analysis display product.

The skew correction and actual interface analysis required approximately 10 hours of effort by the technician; output of the display products was a 2-hour process, requiring the computer operator only. Approximately 3 hours were required by the investigator and the technician to quality check the final product. Table II contains the complete time requirements for this demonstration project, including the analysis of both frames of data.

TABLE II.- MAN-HOURS REQUIRED FOR ANALYSIS OF ALABAMA TIDAL SHORELINE

Team member	Data selection	Classification	Isolation	Scale	Analysis	Total
Investigator	4	5	4	0	3	16
Technician	0	30	18	6	13	67
Computer operator	0	4	18	0	2	24
Geological Survey representative	0	0	3	0	0	3
Total	4	39	43	6	18	110

TECHNIQUE EVALUATION

The interface analysis technique was evaluated with respect to the accuracy and reproducibility of the measurement of natural shoreline features using Landsat data. The results of this study demonstrate that the accurate and reproducible measurement of natural features on Landsat data is possible, and that the measurements made with these data are comparable to measurements made using standard opisometer techniques on 1:24 000-scale maps.

Reproducibility

Seven features in Louisiana, Mississippi, and Alabama were selected for repetitive analysis. All the features were coastal, representing mainland shoreline and offshore islands. The Landsat data including these features were acquired for a minimum of four dates for each feature, classified by using the Water Search procedure previously described, and processed through the interface analysis computer program. The scale factors were determined for each data frame and averaged. At the time this work was performed, it was believed that errors in determining the scale factors usually exceeded the variability of these factors. The scale factor determination procedure has subsequently been improved, and scale factors are now computed and used directly for each individual frame.

Table III contains the results of the analyses. It is clear from this table that the measurements of the same features in different data sets are consistent; the average variation is only 3.48 percent.

The effect of tides on the reproducibility of the shoreline measurements has been studied with this same data set; the resultant correlation was small, 0.26 for all the features. The correlation between tide stage and shoreline length varied among the different features, from -0.11 for Deer Island to -0.04 for the shoreline near Pass Christian, Mississippi, to 0.73 for the shoreline at Long Beach, Mississippi. The latter correlation corresponds to a confidence level of 90 percent. Although apparently linked to tide variation, the changes of shoreline length found for the Long Beach data set were small, having a sample deviation of 4.75 percent and representing a total range of 12 percent.

The tidal effects question was investigated using two other data sets. First, data from two Landsat passes over the Rockefeller Game Refuge in the Louisiana coastal marshlands were analyzed. A decrease of 0.76 kilometer shoreline length at low tide from a total of 87.40 kilometers at high tide was determined, whereas the ratio of land area to water area within the refuge dropped from 0.74 to 0.68 and thus showed a significant decrease of exposed land from low tide to high tide. Second, aircraft data were obtained over another Louisiana marsh area at significantly different tide stages. Photography and multispectral scanner data were acquired during these two different tide conditions from an altitude which permitted acquisition of the data with a minimum mapping unit of less than 10 square meters for the scanner data.

Pairs of frames of the photography representing high- and low-tide conditions in three areas were scaled, registered, and superimposed, and the high- and low-tide shoreline lengths were measured using an opisometer. Figure 11 is one of the areas studied. The yellow area was exposed land at both stages, the black and dark green areas were inundated at both stages, and the red and orange areas were exposed at low tide and inundated at high tide. The area pictured in figure 11 showed an increase in shoreline length of 80 percent from low tide to high tide, whereas one of the other frames showed only a 3-percent increase. The third frame, which could not be registered as well because of some distortion in the original photography, showed a substantial decrease in shoreline length from high tide to low tide.

Analysis of the multispectral scanner data acquired on these two aircraft flights covered a larger area than the photographic analysis. Two areas representing more than 35 square kilometers were processed for MSS imagery as opposed to less than 5 kilometers for the photography. The scanner data first were subjected to the water search and interface analysis in the original form, with the result that shoreline length decreased 17 percent in one area characterized by a tide-level decrease of 23 centimeters and 40 percent in the other area characterized by a tide-level decrease of 15 centimeters. The area which showed the 80-percent increase in the photographic study was included in the second area. The scanner data then were processed to simulate the minimum mapping unit of the Landsat scanner before subjection to the water search and interface analysis. The result was a change with tide of only 2 percent for

Figure 11.- Superposition of aerial photographs acquired at high and low tides. Yellow area was exposed land at both stages; black and dark green areas were inundated at both stages; and red and orange areas were inundated at high tide, exposed land at low tide.

TABLE III.- REPEATABILITY OF LANDSAT ANALYSES

Area	Number of data sets	Mean shoreline length, km	Sample deviation		Correlation with tide stage
			km	percent	
Deer Island	8	15.786	0.313	1.98	-0.11
Broadwater	7	12.676	.343	2.70	.16
Pass Christian	4	22.226	1.047	4.71	-.04
Long Beach	6	4.621	.219	4.75	.73
Half Moon Island	4	12.717	.731	5.75	.14
Petit Bois Island	4	17.155	.336	1.96	.44
Dauphin Island	5	80.861	2.052	2.54	.53
Mean	--	--	--	3.48	--

the first data set and -3.8 percent for the second set. Both of these differences are within the precision of the measurement (which is not as good for aircraft data as it is for Landsat data). Therefore, it is concluded that, at the spatial resolution of the Landsat scanner, tidal effects are not significant for the areas studied.

It may not seem logical that the changes in shoreline configuration that occur with the variations of tide stage should result in negligible changes in the length of the shoreline, but this is indeed the case. In some areas, a falling tide exposes land that previously was under water; in other areas, small streams, and channels, and tidal ponds are left dry. When a pond dries up, shoreline is lost, but when an island emerges, shoreline is gained. When an island with a complex shoreline at high tide is surrounded by a tidal mudflat, that island will lose shoreline as the tide falls because mudflats generally have smooth shorelines. However, experience has shown that in most cases, the change in shoreline at the resolution of the Landsat multispectral scanner is small and of a translational nature. Therefore, the effects of the changing shoreline are to vary the area of exposed land while maintaining a nearly constant shoreline length. Further, the detectable changes in length tend to cancel over a large area because shoreline length increases in some places and decreases in others with the same variation in tide.

The areas studied to determine the effects of tide on shoreline length were in the Louisiana marshlands and along the Mississippi gulf coast. The areas represent very-low-gradient surfaces on which small tide changes cause very wide areas to be either inundated or exposed. The effects of changing tide on the shoreline configuration are probably as great in the Louisiana coastal marshes as in any other area. It is concluded that the effects of tide will not be significant when data from other regions are eventually studied.

Accuracy

The accuracy of a measurement of interface length is a function of several factors. If the depiction of a feature in the source material is accepted as completely accurate, with the exception of errors introduced as a result of finite sampling of the original scene, an absolute accuracy evaluation would include all errors entering into the formulation and processing of the data. These errors include geometric distortion resulting from the manner in which the image is constructed and misclassification of the color data actually acquired by the remote sensor. This misclassification results in the shifting of picture elements from one land cover category to another. Consequently, the accuracy analysis performed is statistical rather than theoretical and represents the comparison of the results of processing Landsat multispectral scanner data with measurements made using traditional methodology.

As a verification that the technique for automatically detecting and measuring interface length was operating correctly, a series of tests was performed in which features of known length were constructed. They were sampled and then analyzed using the interface detection and measurement technique. The results showed agreement with the actual length of the features to 0.6 percent.

It is impossible to define an absolute accuracy of a shoreline length measurement because there is no absolute definition of shoreline. On a very-small-scale map, the shoreline is basically a smooth line, more properly defined as the coastline. On larger scale maps, the complexity of the shoreline increases to include small islands, bays, jutting peninsulas, and intruding estuaries. On the largest scale maps, the irregularities of the shoreline increase, and, to take an extreme case, surface features may be represented at such large scale that scattered rocks and lumps of mud with a few blades of grass protruding from the water surface have their own circumscribing shoreline.

To examine the effects of varying the minimum mapping unit on the shoreline length measurement in a complex marsh area, aircraft multispectral scanner data were collected at an altitude of 1400 meters. The data therefore represented a minimum mapping unit (one pixel) of approximately 3.4 meters on a side. The data were processed to generate new data sets having pixels with twice, four times, and eight times the dimensions of the original picture element. A fifth set of data containing picture elements with dimensions approximating those of the Landsat scanner was also developed. Portions of these data sets were classified using program Water Search and processed with the interface analysis computer program after precise determination of the scale factors.

The results of this analysis (fig. 12) illustrate that the shoreline lengths for the five areas decreased drastically as the mapping unit, or pixel size, increased. The effects were different in each area, an indication that the effect of mapping unit size variation depends on the configuration of the shoreline.

All of the shoreline measurements made in this study are valid, despite the variation. Therefore, an absolute shoreline length cannot be defined

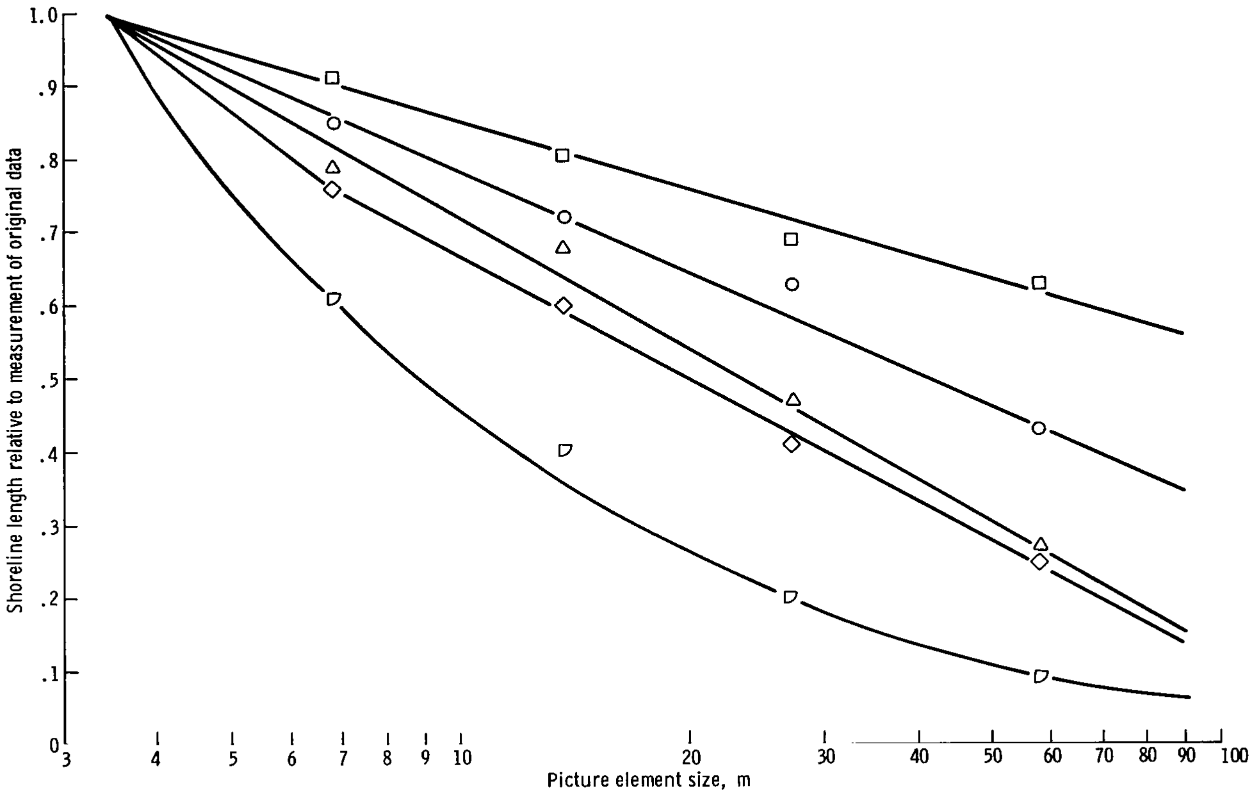


Figure 12.- Effect of minimum mapping unit size on measured shoreline length.

using either maps or remotely acquired data. For a measurement of shoreline length to be of any value, it must have associated with it the spatial resolution or minimum mapping unit size of the source material.

Visual comparison of land/water thematics developed from Landsat data with U.S. Geological Survey USGS 1:250 000- and 1:24 000-scale maps indicated that the larger scale maps represented approximately the same degree of shoreline complexity as the thematics. It should be noted that the minimum mapping unit for the 1:24 000-scale maps, although not rigorously defined, is considerably smaller than the Landsat picture element. Irregularities as small as approximately 3 meters may be shown, and very small streams or even ditches may be shown if they are visible on the aerial photography used to compile the map. A test comparison of actual measurements of shoreline length on the two standard maps and the results of application of the interface analysis technique to the satellite data was performed. The large-scale map measurements agreed very well with the satellite measurement, whereas the small-scale map measurements differed from the other two substantially.

A more detailed comparison was then performed, using the features that were the subject of the reproducibility study and 1:24 000-scale USGS maps. The results of this study are presented in table IV in which the mean values of the repeated satellite measurements are compared with the length of the features as shown on the 1:24 000-scale maps as determined by opisometry, with the restriction that streams less than 30 meters are not measured (as specified in the definition of shoreline used by the National Ocean Survey in their shoreline length determinations). The average magnitude of the deviation between the two sets of measurements was 4.64 percent, with the measurements of Dauphin Island agreeing to 0.61 percent.

TABLE IV.- COMPARISON OF LANDSAT ANALYSES WITH MAPS

Area	Shoreline length, km		Map-Landsat deviation	
	Landsat measurement	Opisometer on maps[a]	km	percent
Deer Island	15.786	14.77	-1.02	6.44
Broadwater	12.676	12.01	-.67	5.25
Pass Christian	22.226	21.04	-1.19	5.34
Long Beach	4.621	4.71	.09	1.93
Half Moon Island	12.717	12.22	-.50	3.91
Petit Bois Island	17.155	15.61	-1.55	9.01
Dauphin Island	80.861	81.36	.51	.61
Mean	--	--	--	4.64

[a]1:24 000-scale USGS quadrangle maps.

CONCLUDING REMARKS

The computer-implemented interface detection and measurement technique described in this report is designed to provide an accurate and inexpensive means of extracting information about interfaces from remotely acquired imagery. The technique consists of a simple 10-step procedure leading to the production of a color-coded or black-and-white display in which the selected interface is clearly defined.

Actual use of the technique for the successful measurement of tidal shoreline length demonstrates an application of interest to geographers and resource managers. Using remotely acquired imagery, the technique could be used to detect and measure the interface between any surface features, natural or manmade, over a broad area.

Application of this technique to repeated Landsat coverage of shoreline features demonstrates the reproducibility of the shoreline length measurement. Tidal variation does not significantly affect the measurement, which has been found to agree with manual measurements made on large-scale maps. The analysis of Landsat data does yield measurements comparable to those made on large-scale maps, but because the minimum mapping unit or pixel size greatly determines the complexity of an interface feature, no interface length measurement should be used without the pixel size associated with it.

The software described in this document is being transferred to the Computer Software Management and Information Center, University of Georgia, Athens, Georgia 30602, through which it will be available to the public. Training courses in the use of the computer programs described in this document, as well as others used by the NASA Earth Resources Laboratory for analysis of satellite data for Earth resources applications, are conducted by the Earth Resources Laboratory in Slidell, Louisiana. Information is available upon request.

Lyndon B. Johnson Space Center
National Aeronautics and Space Administration
Houston, Texas, March 24, 1977
177-55-89-00-72

REFERENCES

1. Whitley, Sidney L.: A Procedure for Automated Land Use Mapping Using Remotely Sensed Multispectral Scanner Data. NASA TR R-434, 1975.

2. Whitley, Sidney L.: Low-Cost Data Analysis Systems for Processing Multispectral Scanner Data. NASA TR R-467, 1976.

3. Procedure Manual for Detection and Location of Surface Water Using ERTS-1 Multispectral Scanner Data: Summary. NASA CR-140215, 1974.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

SPECIAL FOURTH-CLASS RATE
BOOK

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
451

U.S.MAIL

435 001 C1 U E 770610 S00903DS
DEPT OF THE AIR FORCE
AF WEAPONS LABORATORY
ATTN: TECHNICAL LIBRARY (SUL)
KIRTLAND AFB NM 87117

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return

---

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546